Evaluation of shortwave infrared atmospheric correction for ocean color remote sensing of Chesapeake Bay

P. Jeremy Werdell [a,b,*], Bryan A. Franz [a], Sean W. Bailey [a,c]

[a] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[b] Science Systems and Applications, Inc., Lanham, MD 20706, USA
[c] Futuretech Corp., Greenbelt, MD 20770, USA

ABSTRACT

The NASA Moderate Resolution Imaging Spectroradiometer onboard the Aqua platform (MODIS-Aqua) provides a viable data stream for operational water quality monitoring of Chesapeake Bay. Marine geophysical products from MODIS-Aqua depend on the efficacy of the atmospheric correction process, which can be problematic in coastal environments. The operational atmospheric correction algorithm for MODIS-Aqua requires an assumption of negligible near-infrared water-leaving radiance, nL_w(NIR). This assumption progressively degrades with increasing turbidity and, as such, methods exist to account for non-negligible nL_w(NIR) within the atmospheric correction process or to use alternate radiometric bands where the assumption is satisfied, such as those positioned within shortwave infrared (SWIR) region of the spectrum. We evaluated a decade-long time-series of $nL_w(\lambda)$ from MODIS-Aqua in Chesapeake Bay derived using NIR and SWIR bands for atmospheric correction. Low signal-to-noise ratios (SNR) for the SWIR bands of MODIS-Aqua added noise errors to the derived radiances, which produced broad, flat frequency distributions of $nL_w(\lambda)$ relative to those produced using the NIR bands. The SWIR approach produced an increased number of negative $nL_w(\lambda)$ and decreased sample size relative to the NIR approach. Revised vicarious calibration and regional tuning of the scheme to switch between the NIR and SWIR approaches may improve retrievals in Chesapeake Bay, however, poor SNR values for the MODIS-Aqua SWIR bands remain the primary deficiency of the SWIR-based atmospheric correction approach.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The daily synoptic images of Chesapeake Bay provided by satellite ocean color instruments, such as the National Aeronautics and Space Administration (NASA) Moderate Resolution Imaging Spectroradiometer onboard the Aqua platform (MODIS-Aqua), provide viable complementary data streams for operational water quality monitoring activities, which typically rely entirely on *in situ* measurements (Werdell et al., 2009). Briefly, these instruments measure the spectral radiance emanating from top of the atmosphere ($L_t(\lambda)$; μW cm^{-2} nm^{-1} sr^{-1}), at discrete visible and infrared wavelengths. Atmospheric correction algorithms (e.g., Gordon & Wang, 1994) are applied to $L_t(\lambda)$ to remove the contribution of the atmosphere from the total signal and produce estimates of normalized water-leaving radiances ($nL_w(\lambda)$; μW cm^{-2} nm^{-1} sr^{-1}), the light exiting the water mass normalized to a hypothetical condition of an overhead Sun and no atmosphere (Gordon & Clark, 1981). Bio-optical algorithms (e.g., O'Reilly et al., 1998) are applied

to the $nL_w(\lambda)$ to estimate additional marine geophysical properties, including the near-surface concentration of the phytoplankton pigment chlorophyll-a (C_a; mg m^{-3}), a proxy for algal biomass of interest to Chesapeake Bay water quality managers (Chesapeake Bay Program, 1993).

Atmospheric correction algorithms for MODIS-Aqua historically make use of two near-infrared (NIR) bands and an assumption of negligible nL_w (NIR) (≈ 0; commonly referred to as the "black pixel assumption"). The relevant MODIS-Aqua NIR band pair is 748 and 869 nm. While appropriate for the open ocean, the black pixel assumption progressively degrades with increasing marine turbidity, for example, along the coastline and within estuaries. As such, methods exist to account for non-negligible $nL_w(\lambda_{NIR})$ within the atmospheric correction process (Hu et al., 2000; Ruddick et al., 2000; Siegel et al., 2000; Stumpf et al., 2003). Bailey et al. (2010) described the "NIR-correction" approach currently adopted by the NASA Ocean Biology Processing Group (OBPG; McClain et al., 2006), the entity responsible for operationally processing and distributing MODIS-Aqua ocean color products to the research community. This approach builds on Stumpf et al. (2003) and uses a bio-optical model to estimate nL_w(NIR) from nL_w(667).

Wang and Shi (2007) recommended using shortwave infrared (SWIR) bands for atmospheric correction of MODIS-Aqua in areas of

sufficient turbidity that the black pixel assumption fails for $nL_w(\lambda_{NIR})$. More specifically, they proposed using the standard NIR-based atmospheric correction unless a turbidity index (T_{ind}; Shi & Wang, 2007) exceeds a predefined threshold and, thereby, triggers the use of a SWIR-based correction where two SWIR bands are used in lieu of two NIR bands (Wang, 2007). The relevant MODIS-Aqua SWIR band pair is 1240 and 2130 nm (its 1640 nm band does not function properly; Franz et al., 2006). Shi and Wang (2009) concluded that these bands satisfy the black pixel assumption in moderately (1240 nm) to extremely (2130 nm) turbid waters. This combined "NIR–SWIR" atmospheric correction approach has been evaluated extensively in several geographic locations (e.g., the La Plata Estuary and the east coasts of the U.S. and China; (Shi & Wang, 2009; Wang & Shi, 2007)) and vetted against an independent, globally-distributed *in situ* data set (Wang et al., 2009).

Applying the NIR–SWIR approach in Chesapeake Bay conceptually improves the quality of MODIS-Aqua $nL_w(\lambda)$, as the black pixel assumption often fails for $nL_w(\lambda_{NIR})$ in this region (see, e.g., Fig. 3 in Gitelson et al., 2007). Wang and Shi (2007) cautioned, however, that low sensor signal-to-noise (SNR) values in the MODIS-Aqua SWIR bands add noise errors to derived $nL_w(\lambda)$ when these bands are used for atmospheric correction (see also Gordon, 1997; Hu et al., 2001). Wang (2007) recommended that SNR values for 1240 and 2130 nm exceed 148 and 81, respectively (for maritime aerosols with aerosol optical thickness in the NIR of 0.1) to be considered viable for atmospheric correction. True SNR values for MODIS-Aqua SWIR bands are nearly an order of magnitude lower, recently calculated to be 25 and 12, respectively (Franz et al., 2006).

We extend previous work (Werdell et al., 2009) by describing the regional and seasonal variability of MODIS-Aqua $nL_w(\lambda)$ retrievals in Chesapeake Bay derived using the combined NIR–SWIR atmospheric correction approach. While the SWIR atmospheric correction approach has demonstrated advantages in excessively turbid water (Wang & Shi, 2007; Wang et al., 2009), the MODIS-Aqua instrument itself may be inadequate for its use given the low SNR values of the SWIR bands. We first compare MODIS-Aqua $nL_w(\lambda)$ retrievals in Chesapeake Bay derived using the NIR and NIR–SWIR atmospheric correction algorithms as implemented by the OBPG. We then explore how noise introduced by low SNR values propagates into $nL_w(\lambda)$ when the SWIR bands are used for atmospheric correction. Given the paucity of *in situ* radiometry in Chesapeake Bay available for direct comparison, we largely focus on comparing qualitative aspects of $nL_w(\lambda)$, such as percentage of negative retrievals and sample size.

2. Methods

2.1. Data acquisition and processing

We acquired ~6000 MODIS-Aqua spatially extracted Level-1A files containing all or part of the Bay from the OBPG. The time-series spanned June 2002 through December 2008 at ~1 km² nadir spatial resolution. We used the OBPG processing software L2GEN configured for MODIS-Aqua Reprocessing 2009 (OBPG, 2009) to generate Level-2 files containing $nL_w(\lambda)$, $L_t(\lambda)$, spectral aerosol reflectance ($\rho_a(\lambda)$; unitless), C_a, T_{ind}, and spectral aerosol optical thickness ($AOT(\lambda)$; unitless). With this configuration, we generated Level-2 data files that are identical to those distributed by the OBPG to the research community, with the exception that we included additional data products. We calculated C_a using the MODIS "three-band ocean color chlorophyll" algorithm (OC3M; Werdell, 2009). We calculated T_{ind} for each pixel following Eq. (3) of Shi and Wang (2007). For NIR–SWIR processing, we used the 1240 and 2130 nm bands for atmospheric correction (in lieu of the 748 and 869 nm bands) when $T_{ind} \geq 1.3$, as recommended by Wang and Shi (2007). Shi and Wang (2007) and Wang et al. (2009) proposed alternate thresholds for switching from NIR to SWIR processing, which we review in Section 4. We adopted

the operational pixel-masking scheme for MODIS-Aqua, as well as the quality control metrics described in Werdell et al. (2009). After application of the latter, ~13 days of data per month remained for analysis. For several scenes, we calculated SNR values using look-up tables (LUTs) of noise equivalent radiance ($NE\Delta L(\lambda)$; $\mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$) estimated on-orbit by the MODIS Characterization Science Team (e.g., Xiong & Barnes, 2003). We defined $SNR = L_t(\lambda)/NE\Delta L(\lambda)$. The LUTs required focal plane detector number and measured $L_t(\lambda)$ as input.

We acquired *in situ* measurements of C_a from the Chesapeake Bay Program (CBP; Chesapeake Bay Program, 1993), $nL_w(\lambda)$ from the NASA bio-Optical Marine Algorithm Data set (NOMAD; Werdell & Bailey, 2005), and $AOT(870)$ from the NASA Aerosol Robotic Network (AERONET; Holben et al., 2001). The pigment data consisted of discrete fluorometric and spectrophotometric samples from the CBP Water Quality Monitoring Data set that encompassed 49 revisited stations in the mainstem Bay. We considered only near-surface samples (depths ≤ 1 m), as the Bay typically has shallow optical depths, and averaged replicate samples. Additional details regarding treatment of C_a are provided in Werdell et al. (2009). The radiometric data consisted of 523 in-water measurements collected between 1996 and 2006 throughout the mainstem Bay by participants in the NASA Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS; Harding & Magnuson, 2003) and Ocean Biology and Biogeochemistry programs. Additional details regarding calculation of $nL_w(\lambda)$ are provided in Werdell and Bailey (2005). The aerosol data consisted of daily averages of the AERONET Level-2 Quality Assured (calibrated, cloud-screened, and quality controlled) $AOT(\lambda)$ products collected at the Chesapeake Lighthouse (COVE; 37.9°N, −75.7°W) and Smithsonian Environmental Research Center (SERC, 38.9°N, −76.5°W) stations.

2.2. Study area and data stratification

We adopted the spatial and temporal stratification described in Werdell et al. (2009), which was based largely on the methods of Magnuson et al. (2004). Briefly, we determined seasons following Northern Hemisphere equinoxes and solstices, using days of year 80, 172, 266, and 355 to define the transitions of winter–spring, spring–summer, summer–fall, and fall–winter. We assigned 37.6° and 38.6° as the latitudinal boundaries between the Lower–Middle and Middle–Upper Bays (see Fig. 2 in Werdell et al., 2009). This stratification follows a salinity gradient, with the Upper and Lower Bays largely oligohaline and polyhaline, respectively. Bio-optical properties of the Bay also maintain moderate geographic dependency, as the winter–spring freshet from the Susquehanna River regulates the spring phytoplankton bloom (Miller et al., 2006), while other non-algal optically active constituents correlate well with riverine discharge (Harding et al., 2005; Zawada et al., 2007). These features, however, are often obscured by tidal cycles, estuarine circulation, frontal features, and sediment resuspension. We excluded dates when satellite sampling sizes for a given region were less than 200 valid (i.e., unmasked) marine pixels. This eliminated <4% of data for the Lower and Middle Bays, but excluded an additional 1–3 days per month for the Upper Bay.

2.3. Data analysis

We adopted the Level-2 "match-up", frequency distribution, and monthly time-series validation methods presented in Werdell et al. (2009). First, we generated Level-2 match-ups by statistically comparing coincident satellite and *in situ* measurements using the operational configuration of the OBPG satellite data product validation system (Bailey & Werdell, 2006). Specifically: (1) temporal coincidence was defined as ± 3 h; (2) satellite values were the filtered median (via the semi-interquartile range) of all unmasked pixels in a 5×5 pixel box centered on the *in situ* target; and (3) satellite values

were excluded when the coefficient of variation within this box exceeded 0.15. Next, we generated frequency distributions and related statistics for each regional and seasonal subset presented in Section 2.2. Distribution statistics included population medians, semi-interquartile ranges (SIQR; the range covered by values such that 50% of values occur with equal probability on either side of the median), and percentages of negative retrievals ($= 100\% \cdot N_{negative}/N_{total}$, where N is number of pixels). We adopted distribution sampling steps of 0.1 and 0.05 for $nL_w(\lambda)$ and log-transformed C_a, respectively. Finally, we generated regional time-series by calculating the monthly geometric means of all available (unmasked) data. Werdell et al. (2009) provides a comprehensive review of the strengths and weaknesses of each of these validation methods.

3. Results

3.1. $nL_w(\lambda)$ from the NIR and NIR–SWIR approaches

The NIR–SWIR atmospheric correction approach returned a higher frequency of negative $nL_w(\lambda)$ than the NIR approach (Figs. 1 and 2A–C, Table 1). This was true for all $nL_w(\lambda)$ in all three Bay regions for all four seasons. On average, the percentage of negative retrievals increased by 57, 26, and 21% for $nL_w(412)$ and by 507, 242, and 179% for $nL_w(443)$ in the Lower, Middle, and Upper Bays, respectively. The NIR–SWIR approach produced broad distributions of $nL_w(443)$ relative to the NIR approach, indicating a general increase in the dynamic range of retrievals (i.e., general flattening of the frequency distributions). On average, the distribution SIQR increased by 76, 78, and 68% for $nL_w(412)$ and 73, 74, and 56% for $nL_w(443)$ for the Lower, Middle, and Upper Bays, respectively. Decreases in the sample sizes of C_a (N_{C_a}) accompanied increases in the number of negative $nL_w(\lambda)$ (Fig. 2D,

Table 2). On average, N_{C_a} decreased by 5.7, 6.6, and 9.4% in the Lower, Middle, and Upper Bays, respectively. We do not present $nL_w(\lambda)$ match-up results, as all viable radiometric match-up stations had $T_{ind} < 1.3$ (i.e., the SWIR bands were never used in the atmospheric correction process).

3.2. Turbidity index, T_{ind}

We evaluated distributions and time-series of T_{ind} to quantify how frequently the SWIR bands were used in the NIR–SWIR atmospheric correction process. T_{ind} varied seasonally, but not geographically (Fig. 3). The Lower, Middle, and Upper Bays showed regionally consistent, repeated seasonal cycles, as demonstrated in both the frequency distributions (Fig. 3A–D) and monthly time-series (Fig. 3E). T_{ind} in spring and summer were consistent, with an average distribution median, SIQR, and percentage ≥ 1.3 of 1.19, 0.15, and 47%, respectively. The latter statistic ($= 100\% \cdot N_{T_{ind} \geq 1.3}/N_{total}$) provided the frequency at which SWIR bands were used in the NIR–SWIR approach (Fig. 3E). T_{ind} in fall and winter were also consistent, with an average distribution median, SIQR, and percentage ≥ 1.3 of 1.31, 0.26, and 62%, respectively. Overall, the SWIR bands were used for atmospheric correction of MODIS-Aqua 15% more frequently in fall–winter than in spring–summer.

3.3. $nL_w(\lambda)$ from the NIR and SWIR approaches

By design, $nL_w(\lambda)$ from the two atmospheric correction approaches differed only where $T_{ind} \geq 1.3$, that is, when the SWIR bands were used for atmospheric correction in the NIR–SWIR approach. In this section, we consider only these "turbid" pixels and, thus, directly compare the use of NIR and SWIR bands for atmospheric
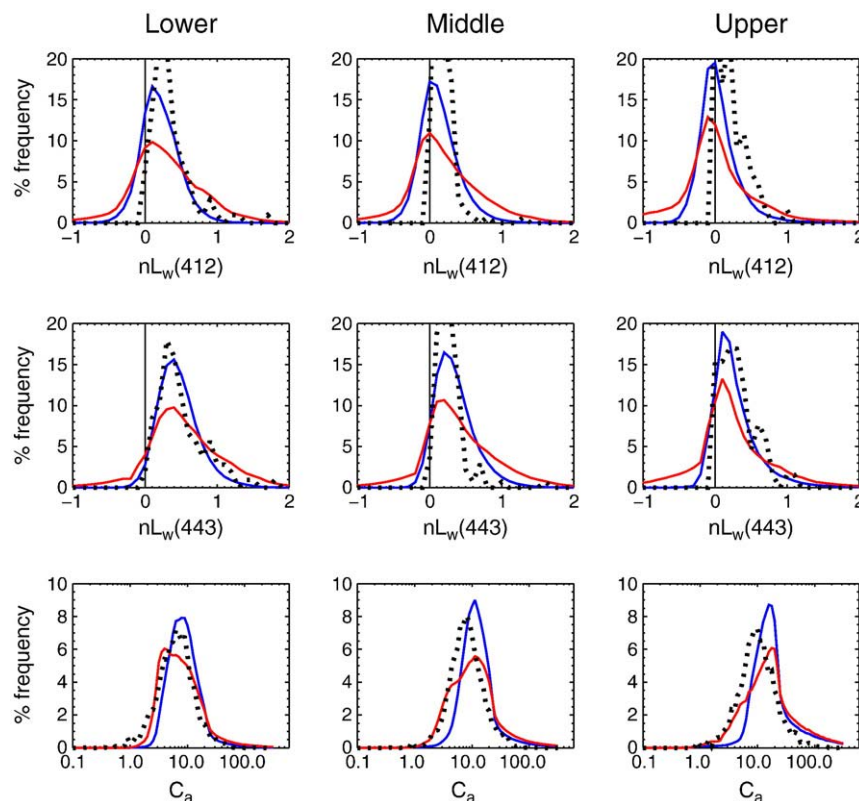

Fig. 1. Frequency distributions of $nL_w(412)$, $nL_w(443)$, and C_a in the Lower, Middle, and Upper Bays. MODIS-Aqua data processed using the NIR and NIR–SWIR atmospheric correction approaches shown as blue and red solid lines, respectively. *In situ* data shown as black dashed lines. Data from all years are included.
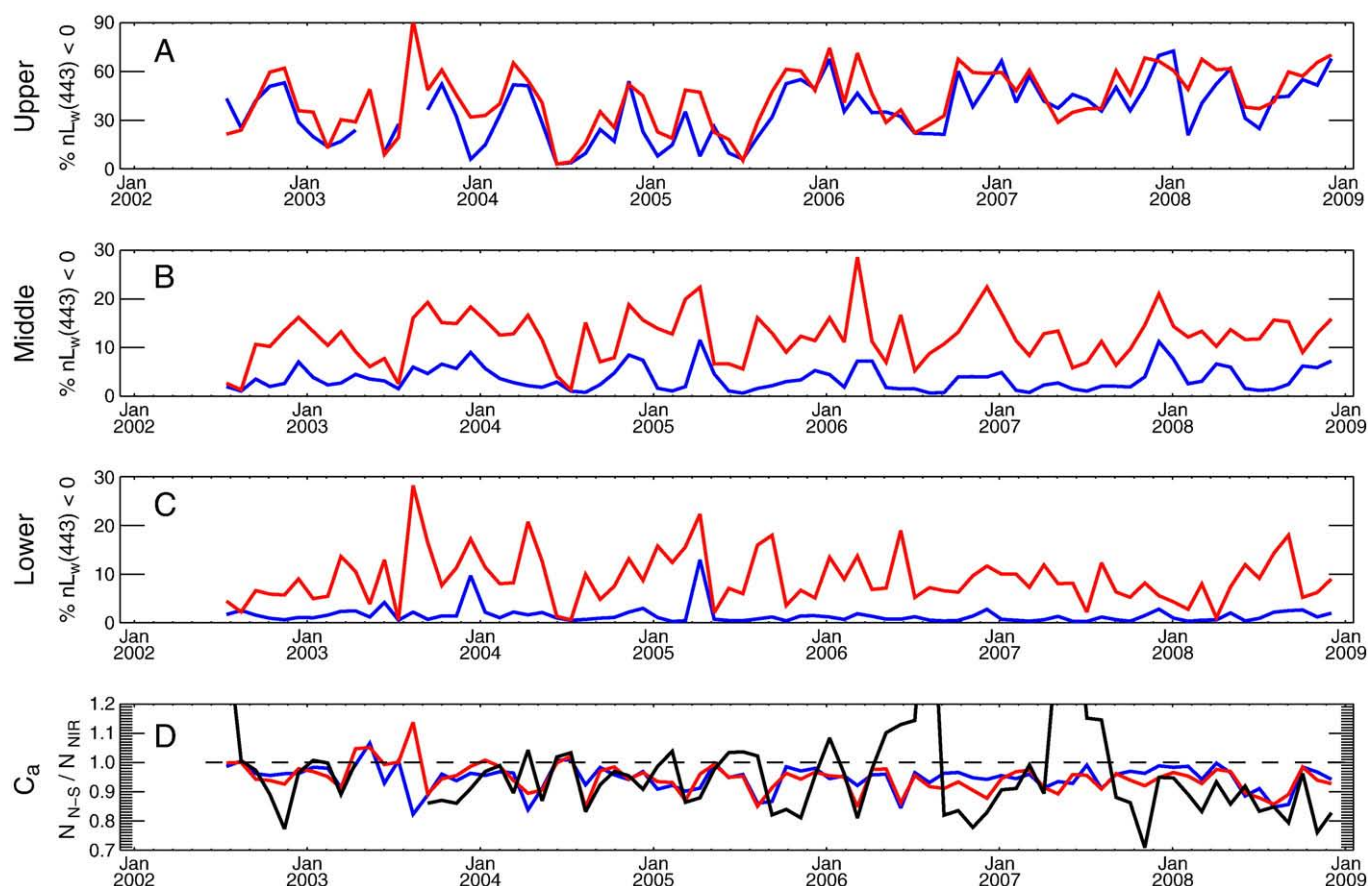
Fig. 2. Monthly averages of the frequency (%) of negative $nL_w(443)$ in the Upper, Middle, and Lower Bays (panels A–C). MODIS-Aqua data processed using the NIR and NIR–SWIR approaches shown as blue and red solid lines, respectively. Monthly averages of the ratio of C_a sample sizes for the NIR (N_{NIR}) and NIR–SWIR (N_{N-S}) approaches (panel D). Ratios for the Lower, Middle, and Upper Bays shown as blue, red, and black solid lines, respectively.

correction of MODIS-Aqua. The SWIR approach returned a higher frequency of negative $nL_w(\lambda)$ than the NIR approach when $T_{ind} \geq 1.3$ (Fig. 4). As in Section 3.1, this was true for all $nL_w(\lambda)$ in all three Bay regions for all four seasons. On average, the percentage of negative

values increased from 10 to 24, from 17 to 27, and from 24 to 38% for $nL_w(412)$ in the Lower, Middle, and Upper Bays, respectively. Likewise, the percentage of negative retrievals increased from 1 to 19, from 2 to 21, and from 3 to 24% for $nL_w(443)$ in the Lower, Middle, and Upper Bays, respectively. The SWIR approach produced broad distributions of $nL_w(\lambda)$ relative to the NIR approach (Table 3). This increase in dynamic range showed spectral dependence, as the SIQR rose by 150, 133, 53, and 12% in the Lower Bay for $nL_w(412)$, $nL_w(443)$, $nL_w(488)$, and $nL_w(547)$, respectively. The SIQR for these bands changed by 162, 144, 51, and 2% in the Middle Bay and 173, 122, 44, and 3% in the Upper Bay.

3.4. C_a from the NIR, NIR–SWIR, and SWIR approaches

The NIR–SWIR approach returned degraded satellite-to-*in situ* match-up statistics for C_a relative to the NIR approach (Fig. 5, Table 4).

Table 1
Percent frequency (%) of $nL_w(\lambda)<0$ in the Lower, Middle, and Upper Bays for MODIS-Aqua processed using the NIR and NIR–SWIR atmospheric correction approaches.

		Lower Bay		Middle Bay		Upper Bay	
		NIR	NIR–SWIR	NIR	NIR–SWIR	NIR	NIR–SWIR
$nL_w(412)$	Spring	16.2	23.6	23.8	28.4	32.6	40.7
	Summer	6.7	15.8	9.0	18.0	33.9	38.2
	Fall	15.0	22.4	27.7	33.9	43.5	52.9
	Winter	16.7	26.1	30.9	36.5	37.9	47.2
	All	14.3	22.5	24.2	30.6	38.7	47.0
$nL_w(443)$	Spring	1.7	10.5	3.6	12.6	6.0	19.0
	Summer	1.2	10.2	2.0	10.6	9.4	19.9
	Fall	1.4	7.8	4.6	13.8	9.9	28.0
	Winter	1.6	9.1	4.0	13.9	7.3	21.9
	All	1.5	9.1	3.8	13.0	8.4	23.4
$nL_w(488)$	Spring	0.1	4.4	0.3	5.5	0.5	7.7
	Summer	0.2	5.4	0.4	5.9	1.7	10.1
	Fall	0	1.5	0.1	3.1	0.4	9.9
	Winter	0	1.7	0	3.3	0.1	6.2
	All	0.1	2.9	0.2	4.1	0.5	8.3
$nL_w(547)$	Spring	0	0.01	0	0.02	0	0.06
	Summer	0	0	0	0.01	0.01	0.01
	Fall	0	0.01	0	0.01	0	0.03
	Winter	0	0.01	0	0.02	0	0.06
	All	0	0.01	0	0.02	0	0.04

Table 2
Cumulative sample sizes of C_a (in 10^5) in the Lower, Middle, and Upper Bays for the full MODIS-Aqua time-series processed using the NIR and NIR–SWIR atmospheric correction approaches.

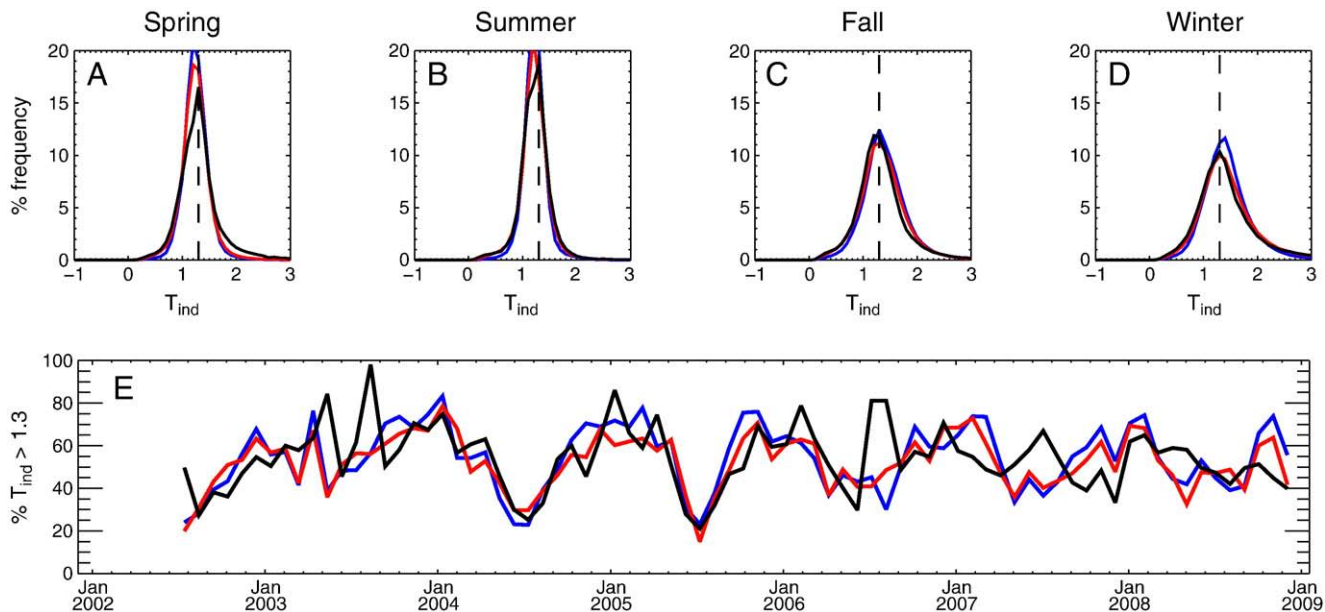|	Lower Bay		Middle Bay		Upper Bay	
	NIR	NIR–SWIR	NIR	NIR–SWIR	NIR	NIR–SWIR
Spring	1.96	1.83	2.58	2.41	0.45	0.43
Summer	1.83	1.68	2.52	2.32	0.32	0.32
Fall	3.38	3.25	4.51	4.24	1.05	0.89
Winter	2.79	2.68	3.67	3.50	0.86	0.81
All	9.97	9.43	13.28	12.46	2.68	2.45

Fig. 3. Seasonal frequency distributions of T_{ind} (panels A–D). $T_{ind} = 1.3$ shown as vertical black dashed lines. Monthly averages of the frequency (%) of $T_{ind} \geq 1.3$ (panel E). Results for the Lower, Middle, and Upper Bays shown as blue, red, and black solid lines, respectively.

The NIR approach returned 5 additional match-up points, however, we excluded these stations and only considered common points in our statistics. The median percent difference (*MPD*) and residual mean squares error (*RMSE*) both increased by 70%, while the median satellite-to-*in situ* ratio decreased 31% from 1.1 to 0.84. While the regression slope improved from 0.83 to 0.95 for the NIR–SWIR approach, the scatter also increased as demonstrated by the elevated *RMSE* and decreased r^2. Of the 71 stations, 52 had $T_{ind} \geq 1.3$ in 17 or more pixels (i.e., $\geq 67\%$ of the 5×5 pixel box). For this subset of stations, the SWIR approach returned degraded satellite-to-*in situ* match-up statistics relative to the NIR approach (Fig. 5, Table 4). The *MPD* and *RMSE* increased by 101 and 56%, respectively, while the
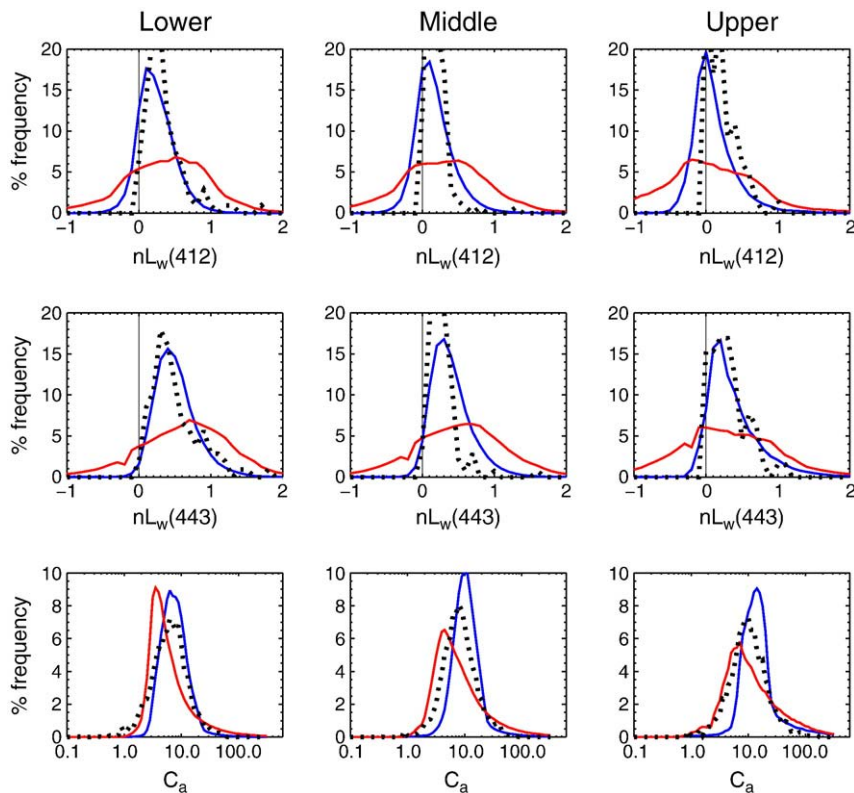

Fig. 4. Frequency distributions of $nL_w(412)$, $nL_w(443)$, and C_a in the Lower, Middle, and Upper Bays. MODIS-Aqua data processed using the NIR and SWIR atmospheric correction approaches shown as blue and red solid lines, respectively. *In situ* data shown as black dashed lines. Data from all years are included. Only pixels with $T_{ind} \geq 1.3$ considered.

Table 3

Medians (M) and semi-interquartile ranges (SIQR) for frequency distributions of $nL_w(\lambda)$ and C_a in the Lower, Middle, and Upper Bays for MODIS-Aqua processed using the NIR and SWIR atmospheric correction approaches. Only stations with $T_{ind} \geq 1.3$ considered.

| | | Lower Bay | | | | Middle Bay | | | | Upper Bay | | | |
| | | NIR | | NIR–SWIR | | NIR | | NIR–SWIR | | NIR | | NIR–SWIR | |
		M	SIQR	M	SIQR	M	SIQR	M	SIQR	M	SIQR	M	SIQR
$nL_w(412)$	Spring	0.14	0.15	0.44	0.42	0.08	0.14	0.36	0.43	0.06	0.18	0.10	0.44
	Summer	0.31	0.15	0.53	0.55	0.28	0.16	0.41	0.48	−0.01	0.13	0.12	0.42
	Fall	0.18	0.17	0.32	0.36	0.07	0.16	0.21	0.38	−0.03	0.14	−0.13	0.42
	Winter	0.10	0.13	0.27	0.39	0.02	0.13	0.15	0.40	0.02	0.18	−0.01	0.47
	All	0.16	0.16	0.35	0.40	0.08	0.16	0.24	0.41	−0.00	0.16	−0.03	0.44
$nL_w(443)$	Spring	0.39	0.16	0.69	0.44	0.30	0.16	0.59	0.45	0.40	0.25	0.42	0.46
	Summer	0.47	0.15	0.71	0.59	0.42	0.17	0.57	0.50	0.16	0.13	0.30	0.44
	Fall	0.45	0.21	0.59	0.38	0.29	0.18	0.43	0.38	0.18	0.17	0.09	0.45
	Winter	0.39	0.17	0.58	0.40	0.27	0.15	0.42	0.41	0.28	0.23	0.27	0.48
	All	0.42	0.18	0.61	0.42	0.30	0.17	0.47	0.42	0.24	0.21	0.23	0.46
$nL_w(488)$	Spring	0.70	0.22	0.95	0.42	0.55	0.25	0.83	0.41	0.86	0.41	0.78	0.44
	Summer	0.70	0.21	0.93	0.55	0.62	0.23	0.78	0.47	0.37	0.18	0.50	0.40
	Fall	0.86	0.30	0.97	0.40	0.62	0.29	0.75	0.38	0.46	0.26	0.43	0.42
	Winter	0.82	0.26	0.97	0.40	0.61	0.26	0.80	0.41	0.59	0.33	0.63	0.47
	All	0.79	0.27	0.96	0.41	0.60	0.27	0.78	0.40	0.54	0.31	0.57	0.45
$nL_w(547)$	Spring	1.33	0.30	1.48	0.38	1.15	0.42	1.36	0.41	1.58	0.57	1.52	0.49
	Summer	1.26	0.26	1.40	0.43	1.22	0.39	1.35	0.44	0.95	0.30	1.06	0.36
	Fall	1.37	0.38	1.47	0.40	1.12	0.44	1.25	0.43	1.00	0.40	1.06	0.43
	Winter	1.44	0.38	1.55	0.41	1.19	0.45	1.37	0.46	1.20	0.47	1.30	0.48
	All	1.36	0.36	1.49	0.40	1.16	0.43	1.32	0.44	1.13	0.45	1.21	0.47
C_a	Spring	9.15	0.32	4.76	0.39	12.30	0.30	5.71	0.47	13.54	0.32	9.16	0.56
	Summer	7.99	0.33	4.35	0.42	10.46	0.28	5.89	0.56	16.95	0.31	7.45	0.58
	Fall	5.56	0.28	4.25	0.41	7.88	0.30	5.35	0.56	12.61	0.32	8.47	0.69
	Winter	7.50	0.33	4.96	0.50	10.10	0.28	6.69	0.61	11.80	0.37	8.35	0.69
	All	6.87	0.35	4.54	0.43	9.64	0.31	5.87	0.57	12.77	0.34	8.46	0.66

median satellite-to-*in situ* ratio decreased 64% from 0.97 to 0.59. The regression slopes were similar for both, but the scatter increased for the SWIR approach as demonstrated by the elevated *RMSE* and decreased r^2.

The SWIR atmospheric approach produced broad distributions of C_a relative to the NIR approach with lower median values (Figs. 1 and 4, Table 3). Considering only pixels with $T_{ind} \geq 1.3$, the SIQR for C_a increased on average by 26, 82, and 92%, while the median decreased by 340, 39, and 34% for the Lower, Middle, and Upper Bays, respectively. These statistics corroborate the satellite-to-*in situ* match-up results, namely the increased dispersion of the match-ups (demonstrated by the *MPD*, *RMSE* and r^2) and decreased satellite-to-*in situ* ratio (Table 4). N_{C_a} decreased when the SWIR approach was used. On average, N_{C_a} decreased by 15, 15, and 4% in the Lower, Middle, and Upper Bays, respectively. The monthly time-series reiterated the tendency of the SWIR approach to produce lower C_a than the NIR approach (Fig. 6). Considering only pixels with $T_{ind} \geq 1.3$, monthly C_a from the NIR approach exceeded those from the SWIR approach by 26,

30, and 32% on average in the Lower, Middle, and Upper Bays, respectively.

4. Discussion

Using SWIR bands for atmospheric correction of MODIS-Aqua produced $nL_w(\lambda)$ with increased dynamic ranges relative to those derived using NIR bands (Figs. 1 and 4, Table 3). The spectral dependency of these increases ultimately resulted in wider distributions of C_a with lower average values (Figs. 4 and 5, Tables 3 and 4). OC3M describes the polynomial best fit that relates C_a to ratios of blue-to-green radiances (Werdell, 2009); therefore, differences in C_a are predictable based on observed differences in $nL_w(\lambda)$. The ratio of $nL_w(488)$ to $nL_w(547)$ increased in dynamic range for the SWIR atmospheric correction approach relative to the NIR approach, which led to the observed increase in the dynamic range of C_a. Shifts in the average values of these radiance ratios led to decreased average C_a. In the Lower Bay, for example, the elevated $nL_w(488)$ to $nL_w(547)$ ratio
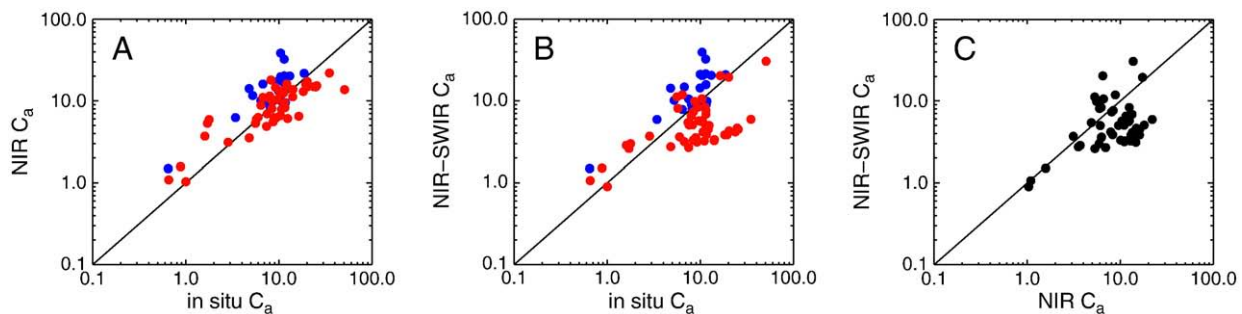

Fig. 5. MODIS-Aqua satellite-to-*in situ* C_a validation ("match-up") results for the NIR and NIR–SWIR atmospheric correction approaches (panels A and B). Stations with $T_{ind} \geq 1.3$ shown in red ($N = 52$). Stations with $T_{ind} < 1.3$ shown in blue ($N = 19$). Comparison of MODIS-Aqua C_a from the two approaches for common stations with $T_{ind} \geq 1.3$ (panel C).

Table 4
C_a satellite-to-*in situ* validation ("match-up") statistics for MODIS-Aqua data processed using the NIR and NIR–SWIR atmospheric correction approaches. Total sample size is 71.

		MPD	Ratio	r^2	b ($\pm$SE)	a ($\pm$SE)	RMSE
NIR vs.	All	33.5	1.10	0.62	0.83 ($\pm$0.06)	0.21 ($\pm$0.06)	0.20
in situ	$T_{ind} \geq 1.3$	28.8	0.97	0.72	0.75 ($\pm$0.06)	0.23 ($\pm$0.06)	0.16
NIR–SWIR	All	57.1	0.84	0.28	0.95 ($\pm$0.11)	-0.06 ($\pm$0.11)	0.34
vs. *in situ*	$T_{shi} \geq 1.3$	58.0	0.59	0.37	0.73 ($\pm$0.10)	0.01 ($\pm$0.09)	0.25
NIR vs. NIR–SWIR	$T_{ind} \geq 1.3$	53.6	0.53	0.22	0.98 ($\pm$0.15)	-0.21 ($\pm$0.14)	0.30

from the SWIR approach $(0.96/1.49 = 0.64)$ relative to the NIR approach $(0.79/1.36 = 0.58)$ resulted in a depressed median C_a (4.54 vs. 6.87 mg m^{-3}; Table 3).

The relative performance of the two atmospheric correction approaches varied little with changing in-water and atmospheric conditions of Chesapeake Bay (Fig. 6). Residuals in the monthly time-series of C_a $(= C_a^{MODIS} - C_a^{insitu})$ showed minor sensitivity to marine turbidity (via T_{ind}), with simple correlation coefficients (r) less than ± 0.28 for both approaches in all regions but one (-0.48 for the NIR approach in the Lower Bay). Restated, biases in satellite-derived C_a from both approaches varied weakly with T_{ind} (which, to a first order, eliminates the need to raise T_{ind} beyond 1.3). Similarly, residuals in the monthly time-series of AOT $(= AOT^{MODIS} - AOT^{insitu})$ showed little sensitivity to T_{ind}, with $0.22 < r < 0.55$ for both approaches in all regions. The modest correlation ($r \sim 0.5$) results from both atmo-

spheric correction approaches underestimating $AOT(865)$ in summer, where T_{ind} is low relative to fall and winter. $AOT(865)$ from the two atmospheric correction approaches showed moderate correlation with r^2 and Type I linear regression slopes (NIR vs. SWIR) of 0.26 and 1.8, 0.54 and 1.7, and 0.64 and 1.7 for the Upper, Middle, and Lower Bays, respectively.

With the exception of the Level-2 satellite-to-*in situ* match-ups, we felt that qualitative skill assessment of the varied atmospheric correction approaches using field measurements was somewhat prohibitive in this study. First, *in situ* $nL_w(\lambda)$ were too scarce to enable meaningful statistical comparisons in the frequency distributions and monthly time-series. Second, *in situ* C_a were not easily subsampled to isolate stations with $T_{ind} \geq 1.3$, which is largely a radiometric measurement. Third, Werdell et al. (2009) pointed out deficiencies in the globally-parameterized OC3M for use in Chesapeake Bay. Regional algorithms tuned to either atmospheric correction approach are easily developed, such as the OC_{corr} and OC_{sat} algorithms described in Werdell et al. (2009). Finally, different analyses (e.g., the Level-2 satellite-to-*in situ* match-ups, frequency distributions, and time-series) present data product validation statistics in different ways, with meaningful interpretation of these statistics depending on the questions being posed.

For example, consider that C_a from MODIS-Aqua has been shown to be elevated compared to *in situ* measurements (e.g., Werdell et al., 2009). The NIR–SWIR approach showed improved ratios of monthly C_a^{MODIS} to C_a^{insitu} relative to the NIR approach in the all three regions,
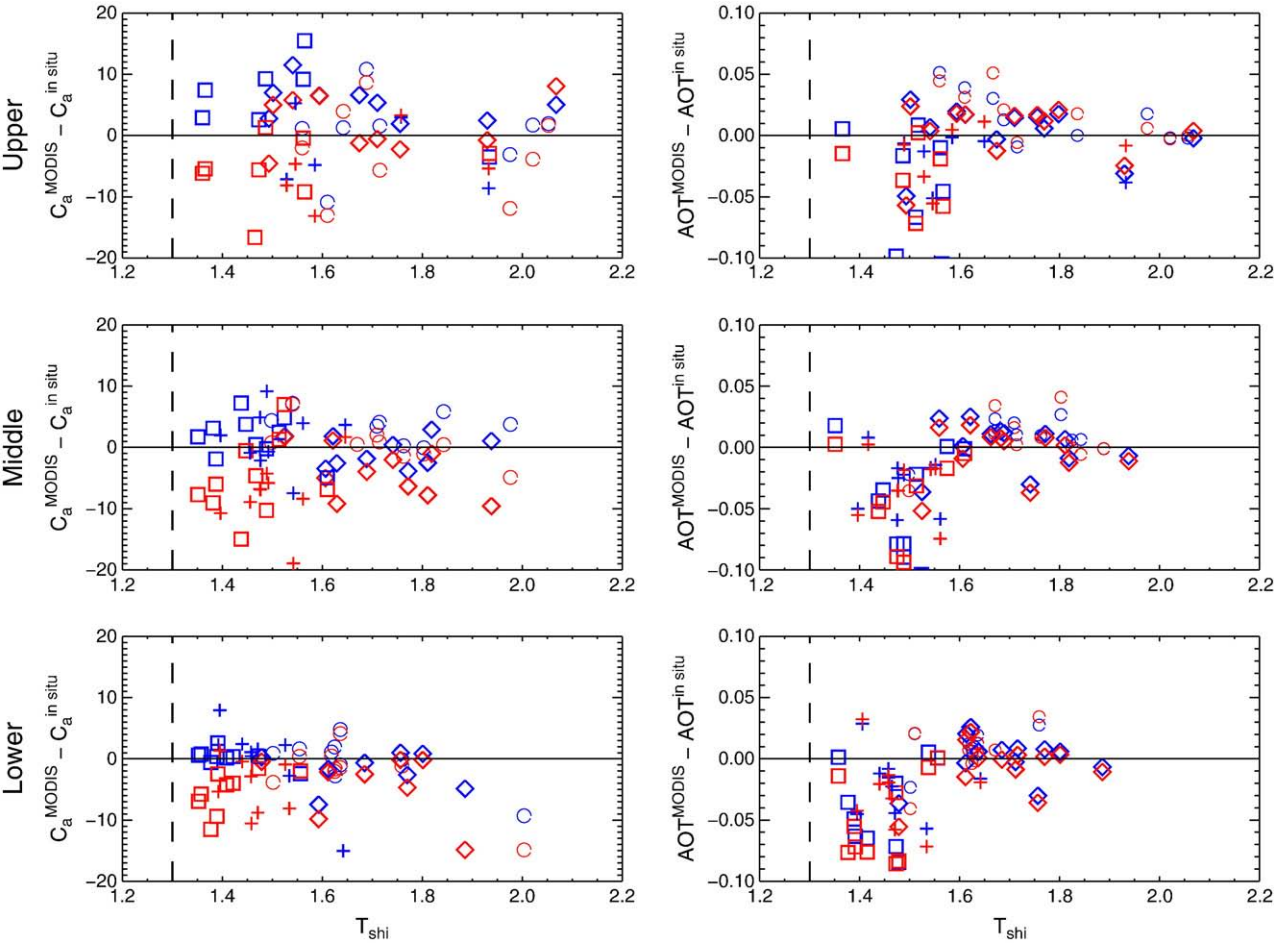

Fig. 6. Residuals from the monthly time-series shown as a function of T_{ind} for the Upper, Middle, and Lower Bays. Residuals from the C_a $(= C_a^{MODIS} - C_a^{insitu})$ and $AOT(865)$ $(= AOT(865)^{MODIS} - AOT(865)^{insitu})$ time-series shown in the left and right columns, respectively. MODIS-Aqua data processed using the NIR and SWIR approaches shown as blue and red symbols, respectively. Data collected in spring, summer, fall, and winter shown as crosses, squares, diamonds, and circles, respectively.

which is in contrast to the Level-2 match-ups results (Table 4). The mean ratios for the NIR and NIR–SWIR approaches were 1.21 and 1.11 in the Lower Bay, 1.32 and 1.27 in the Middle Bay, and 1.79 and 1.47 in the Upper Bay. The monthly time-series for the three regions show C_a from the NIR–SWIR approach to be 14% lower than C_a from the NIR approach when all pixels are considered. The percentage rises to 30% when only pixels with $T_{ind} \geq 1.3$ are considered. But, this reduction in average C_a carried a "cost" of increasing the percentage of negative $nL_w(412)$ by 140, 58, and 58% in the Lower, Middle, and Upper Bays when only pixels with $T_{ind} \geq 1.3$ are considered. The NIR–SWIR atmospheric correction approach also returned fewer valid C_a retrievals (Table 2).

Ultimately, using SWIR bands for atmospheric correction produced $nL_w(\lambda)$ frequency distributions with broad, flat shapes relative to those produced using NIR bands (Fig. 4, Table 3), which supports our hypothesis that poor SNR in these bands adds noise errors to the derived radiances (see also Gordon, 1997; Hu et al., 2001). An example MODIS-Aqua transect of the mainstem Chesapeake Bay showed mean SNR for its 748, 869, 1240, and 2130 nm bands to be 1410, 1119, 52, and 29, respectively (Fig. 7). We verified these values by repeating this exercise for several scenes that spanned different seasons and years. Wang (2007) suggested that SNR for the 1240 and 2130 nm bands be nearly three times as large prior to their use in atmospheric correction. Potentially, the SNRs for the SWIR bands could be increased through spatial smoothing (e.g., by replacing the 1 km^2 pixel with the average of the surrounding 3×3 or 10×10 pixel box), based on the idea that atmospheric aerosols vary on larger spatial scales. Such averaging, however, can introduce straylight and cloud adjacency effects into the aerosol reflectance, and thus dilate the impact of such artifacts on the visible-band radiometric retrievals. In narrow, inland bodies of water such as Chesapeake Bay, straylight from adjacent land would be especially problematic.

Noise in $L_t(\lambda_{NIR})$ or $L_t(\lambda_{SWIR})$ propagates into the aerosol reflectances calculated as part of the atmospheric correction process. Comparison of $\rho_a(\lambda)$ from the NIR and SWIR approaches for pixels with $T_{ind} \geq 1.3$ showed that differences in aerosol reflectances contribute significantly to differences in $nL_w(\lambda)$ (Fig. 8). Frequency distributions for the SWIR approach showed slightly elevated and significantly reduced $\rho_a(443)$ relative to those for the NIR approach (Fig. 8A–C). The peak near $\rho_a(443) = 0$ in the frequency distributions for the SWIR approach indicates that either the retrieved atmospheric signal in the 2130 nm band is often anomalously low or the relative signal between the 1240 and 2130 nm bands is anomalously low. Analysis of the Angstrom exponent retrievals (not shown) suggests the latter; when the signal levels are low, the SWIR approach cannot retrieve the correct aerosol spectral dependence and tends toward spectrally flat models, thus underestimating the aerosol reflectance in the shorter wavelengths. In fact, at low signals the SWIR approach often predicts aerosol spectral dependence that is beyond the range of the current aerosol models, as it would be associated with unrealistically large aerosol size distributions. Conversely, the SWIR approach tends to overestimate aerosol reflectance at higher aerosol loads, thus resulting in negative $nL_w(\lambda)$ (Fig. 8D). This problem could be mitigated through vicarious calibration, by decreasing the calibration of the 2130 nm bands while maintaining or decreasing the spectral slope between the 1240 and 2130 nm bands, but that would exacerbate the underestimation problem at low aerosol loads. Another possible reason for the overestimation of aerosol reflectance in the SWIR is simply that, in the highly reflective waters of Chesapeake Bay, $nL_w(1240)$ is still significant (see Shi & Wang, 2009). The SWIR approach would treat any such residual water-leaving radiance as aerosol, thus over estimating the spectral slope and over-subtracting the aerosols in the visible wavelengths.

L2GEN differs somewhat from the Level-2 data processor used in Wang and Shi (2007) and Wang et al. (2009). To the best of our knowledge, however, L2GEN uses sufficiently similar LUTs within the atmospheric correction process (including aerosol tables with polarization sensitivity), cloud masking, and vicarious calibration of the visible and NIR bands (Franz et al., 2007) to produce comparable results (nevertheless, OBPG products serve as the primary resource of the research community). Poor SNR inhibited robust vicarious calibration of the MODIS-Aqua SWIR bands. An attempt was made to vicariously calibrate the SWIR bands using the approach of Franz et al. (2007) with the assumption of $nL_w(SWIR) = 0$ at the Marine Optical Buoy calibration site. The resulting gains for 1240 and 2130 nm were 1.0474 and 1.1563, respectively. The variance of these gains, however, significantly exceeded the mean values, such that they were statistically consistent with unity. As such, the OBPG applies gains of unity to both bands. As a sensitivity study, we reprocessed the MODIS-Aqua Chesapeake Bay time-series using vicarious gains of 1.0474 and 1.1563 for the 1240 and 2130 nm bands. We reexecuted each analysis, but found all results to be largely identical to those presented above. Changing the gains increased the aerosol reflectance at 2130 nm, but decreased the spectral slope, and thus, yielded virtually no change in the retrieved $nL_w(\lambda)$. While, to a first order, this eliminates vicarious calibration of the SWIR as a mechanism for the degraded $nL_w(\lambda)$ retrievals, we encourage conscientious reanalysis of calibration in future work.

As a second sensitivity study, we reprocessed the MODIS-Aqua Chesapeake Bay time-series using the NIR-to-SWIR switching mechanism proposed by Wang et al. (2009). Rather than switch to the SWIR approach when $T_{ind} \geq 1.3$, this alternative scheme switched when $T_{ind} \geq 1.05$, $C_a > 1$ mg m^{-3}, and $nL_w(869) \geq 0.08$ μW cm^{-2} nm^{-1} sr^{-1}, with C_a and $nL_w(869)$ calculated using the SWIR approach. In Chesapeake Bay, T_{ind} and C_a routinely exceeded 1.05 and 1 mg m^{-3}, respectively (Figs. 1 and 3), such that $nL_w(869)$ predominantly controlled switching between the NIR and SWIR approaches. We reexecuted each analysis, but found all results to be qualitatively identical to those presented above. Changing the switching mechanism reduced the frequency with which the SWIR approach was used by 9, 26, and 63% in the Lower, Middle, and Upper Bays, respectively,
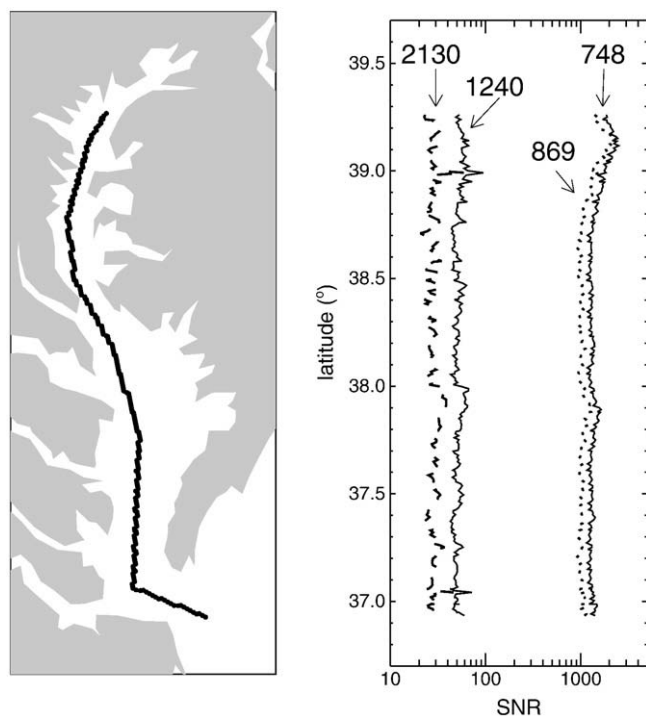

Fig. 7. Signal-to-noise (SNR) ratios for the 748, 869, 1240, and 2130 bands of MODIS-Aqua for a transect through the Chesapeake Bay mainstem. The MODIS-Aqua scene presented was collected on 5 April 2004.
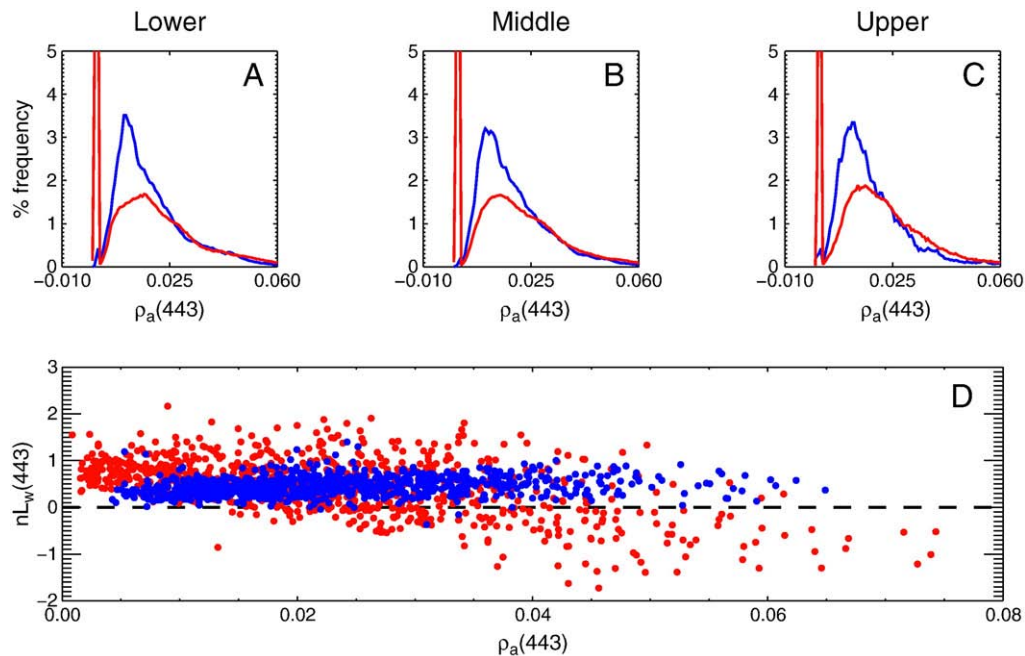
Fig. 8. Frequency distributions of $\rho_a(443)$ in the Lower, Middle, and Upper Bays (panels A–C). $nL_w(443)$ as a function of $\rho_a(443)$ (panel D). MODIS-Aqua data processed using the NIR and SWIR atmospheric correction approaches shown as blue and red solid lines or circles, respectively. Only pixels with $T_{ind} \geq 1.3$ considered.

but frequency distributions of $nL_w(\lambda)$ continued to show broad dynamic ranges and increased numbers of negative retrievals (Fig. 9). Further evaluation of the accuracy of T_{ind} or robustness of the switching schemes in Chesapeake Bay is beyond the scope of this work. In general, however, we recommend the development of a regionally tuned switching mechanism for future efforts.

5. Conclusions

In Chesapeake Bay, ocean color atmospheric correction approaches must account for $nL_w(NIR)$, as water-leaving radiances are rarely negligible in this part of the spectrum (that is, the black pixel assumption often fails). Use of the SWIR bands for atmospheric
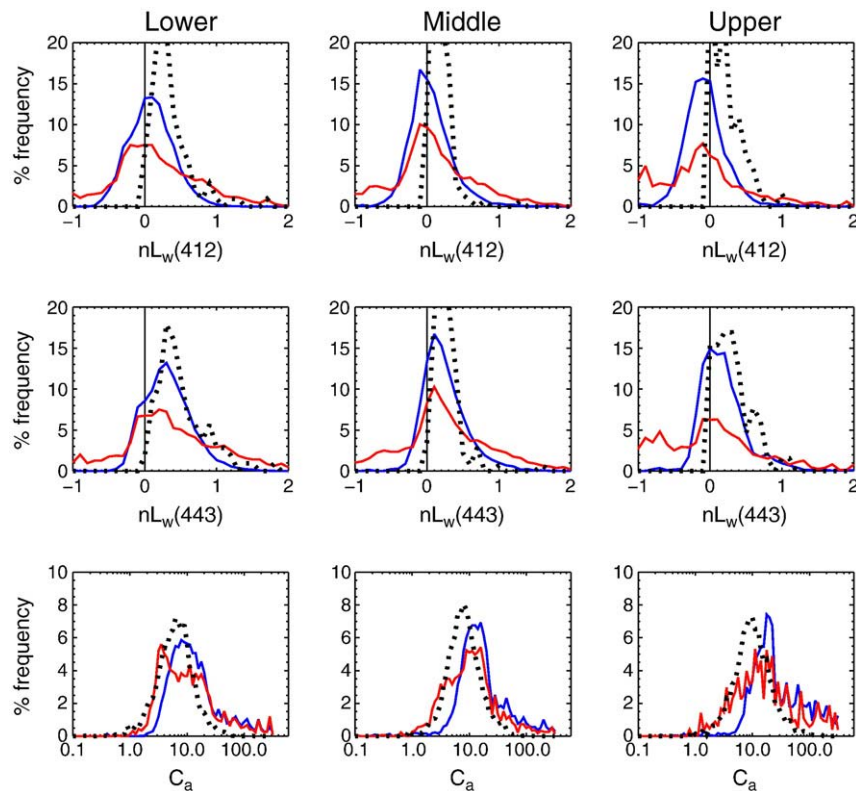

Fig. 9. Frequency distributions of $nL_w(412)$, $nL_w(443)$, and C_a in the Lower, Middle, and Upper Bays. MODIS-Aqua data processed using the NIR and SWIR atmospheric correction approaches shown as blue and red solid lines, respectively. *In situ* data shown as black dashed lines. Data from all years are included. Only pixels with $T_{ind} \geq 1.05$, $C_a \geq 1$ mg m^{-3}, and $nL_w(869) \geq 0.08$ μW cm^{-2} nm^{-1} sr^{-1} considered.

correction provides an attractive alternative for MODIS-Aqua data processing, as its 1240 and 2130 nm bands satisfy the black pixel assumption in moderately (1240 nm) to extremely (2130 nm) turbid waters (Shi & Wang, 2009). Unfortunately, the MODIS-Aqua instrument appears inadequate for SWIR-based atmospheric correction because of prohibitively low SNR values for these bands. The SNR values of 1240 and 2130 nm bands are nearly an order of magnitude lower than simulations suggest are required (Franz et al., 2006; Wang, 2007). Using SWIR bands for atmospheric correction of MODIS-Aqua produced frequency distributions of $nL_w(\lambda)$ with broad, flat shapes relative to those produced using NIR bands, which likely resulted from poor SNR in the SWIR bands adding noise errors to the derived radiances. Spectral differences in $nL_w(\lambda)$ produced lower C_a for the SWIR approach relative to the NIR approach (often desirable in Chesapeake Bay when using OC3M), but with the cost of increased numbers of negative $nL_w(\lambda)$ and decreased sample sizes. While additional attention to vicarious calibration and tuning of the scheme to switch between the NIR and SWIR approaches may improve retrievals somewhat in Chesapeake Bay, poor SNR values for the MODIS-Aqua SWIR bands persist as the principal deficiency of this alternative atmospheric correction approach. Our analyses provide a cautionary tale for future satellite instruments (in particular, those with similar or lesser SNRs) whose proposed data processing includes use of SWIR bands.

Acknowledgements

References

Bailey, S. W., Franz, B. A., & Werdell, P. J. (2010). Updated NIR water-leaving radiance estimation for ocean color data processing. *Optics Express*, *18*, 7521−7527.
Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, *102*, 12−23.
Chesapeake Bay Program (1993). *Guide to Using Chesapeake Bay Program Water Quality Monitoring Data. CBP/TRS 78/92, Chesapeake Bay Program, Annapolis, Maryland.*
Franz, B. A., Bailey, S. W., Werdell, P. J., & McClain, C. R. (2007). Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. *Applied Optics*, *46*, 5068−5082.
Franz, B. A., Werdell, P. J., Meister, G., Kwiatkowska, E. J., Bailey, S. W., Ahmad, Z., et al. (2006). MODIS land bands for ocean remote sensing applications. *Proceedings of Ocean Optics XVIII, Montreal, Canada*.
Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*, 464−472.
Gordon, H. R. (1997). Atmospheric correction of ocean color imagery in the Earth Observing System era. *Journal of Geophysical Research*, *102*, 17081−17106.
Gordon, H. R., & Clark, D. K. (1981). Clear water radiances for atmospheric correction of coastal zone color scanner imagery. *Applied Optics*, *20*, 4175−4180.
Gordon, H. R., & Wang, M. (1994). Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. *Applied Optics*, *33*, 443−452.

Harding, L. W., Jr., & Magnuson, A. (2003). Bio-optical and remote sensing observations in Chesapeake Bay. In G. S. Fargion & C.R. McClain (Eds.), *SIMBIOS Project 2003 Annual Report. NASA Tech. Memo. 212251, NASA Goddard Space Flight Center, Greenbelt, Maryland* (pp. 84−97)..
Harding, L. W., Jr., Magnuson, A., & Mallonee, M. E. (2005). SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic Bight. *Estuarine, Coastal and Shelf Science*, *62*, 75−94.
Holben, B. N., Tanré, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., et al. (2001). An emerging ground-based aerosol climatology: aerosol optical depth from AERONET. *Journal of Geophysical Research*, *106*, 12067−12097.
Hu, C., Carder, K. L., & Muller-Karger, F. E. (2000). Atmospheric correction of SeaWiFS imagery over turbid coastal waters: A practical method. *Remote Sensing of Environment*, *74*, 195−206.
Hu, C., Carder, K. L., & Muller-Karger, F. E. (2001). How precise are SeaWiFS ocean color estimates? Implications of digitization noise errors. *Remote Sensing of Environment*, *76*, 239−249.
Magnuson, A., Harding, L. W., Jr., Mallonee, M. E., & Adolf, J. (2004). Bio-optical model for Chesapeake Bay and the Middle Atlantic Bight. *Estuarine, Coastal and Shelf Science*, *61*, 403−424.
McClain, C. R., Hooker, S. B., Feldman, G. C., & Bontempi, P. (2006). Satellite data for ocean biology, biogeochemistry, and climate research. *EOS Transactions AGU*, *87* (34), 337−343.
Miller, W. D., Kimmel, D., & Harding, L. W., Jr. (2006). Predicting spring discharge of the Susquehanna River from a winter synoptic climatology for the eastern United States. *Water Resources Research*, *42*. doi:10.1029/2005WR004270
OBPG (2009). Ocean Color Reprocessing 2009. (URL). http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/
O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, *103*, 24937−24953.
Ruddick, K. G., Ovidio, F., & Rijkeboer, M. (2000). Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Applied Optics*, *39*, 897−912.
Shi, W., & Wang, M. (2007). Detection of turbid waters and absorbing aerosols for the MODIS ocean color data processing. *Remote Sensing of Environment*, *110*, 149−161.
Shi, W., & Wang, M. (2009). An assessment of the black pixel assumption for MODIS SWIR bands. *Remote Sensing of Environment*, *113*, 1587−1597.
Siegel, D. A., Wang, M., Maritorena, S., & Robinson, W. (2000). Atmospheric correction of satellite ocean color imagery: The black pixel assumption. *Applied Optics*, *39*, 3582−3591.
Stumpf, R. P., Arnone, R. A., Gould, R. W., Jr., Martinolich, P. M., & Ransibrahmanakul, V. (2003). A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from seawifs in coastal waters. *NASA Tech. Memo. 206892, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD*.
Wang, M. (2007). Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: Simulations. *Applied Optics*, *46*, 1535−1547.
Wang, M., & Shi, W. (2007). The NIR–SWIR combined atmospheric correction approach for MODIS ocean color data processing. *Optics Express*, *15*, 15722−15733.
Wang, M., Son, S., & Shi, W. (2009). Evaluation of MODIS SWIR and NIR–SWIR atmospheric correction algorithms using SeaBASS data. *Remote Sensing of Environment*, *113*, 635−644.
Werdell, P. J. (2009). Ocean Color Chlorophyll (OC) v6. (URL). http://oceancolor.gsfc.nasa.gov/ANALYSIS/ocv6/
Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, *98*, 122−140.
Werdell, P. J., Bailey, S. W., Franz, B. A., Harding, L. W., Jr., Feldman, G. C., & McClain, C. R. (2009). Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua. *Remote Sensing of Environment*, *113*, 1319−1330.
Xiong, X., & Barnes, W. (2003). Early on-orbit calibration results from Aqua MODIS. *Proceedings of SPIE*, *4881*, 327−336.
Zawada, D. G., Hu, C., Clayton, T., Chen, Z., Brock, J. C., & Muller-Karger, F. E. (2007). Remote sensing of particle backscattering in Chesapeake Bay: A 6-year SeaWiFS retrospective view. *Estuarine, Coastal and Shelf Science*, *73*, 792−806.